Article:

Thornton, Michael Alan orcid.org/0000-0002-4470-809X and Chambers, Marcus (2016) The Exact Discretisation of CARMA Models with Applications in Finance. Journal of empirical finance. 739 - 761. ISSN 0927-5398

https://doi.org/10.1016/j.jempfin.2016.03.006

The Exact Discretisation of CARMA Models with Applications in Finance

Michael A. Thornton
University of York

and

Marcus J. Chambers
University of Essex

March 2016

Abstract

The problem of estimating a continuous time model using discretely observed data is common in empirical finance. This paper uses recently developed methods of deriving the exact discrete representation for a continuous time ARMA (autoregressive moving average) system of order p, q to consider three popular models in finance. Our results for two benchmark term structure models show that higher order ARMA processes provide a significantly better fit than standard Ornstein-Uhlenbeck processes. We then explore present value models linking stock prices and dividends in the presence of cointegration. Our methods enable us to take account of the fact that the two variables are observed in fundamentally different ways by explicitly modelling the data as mixed stock-flow type, which we then compare with the (more common, but incorrect) treatment of dividends as a stock variable.

Keywords. Continuous time ARMA process; discrete time representation; present value; term structure.

J.E.L. classification number. C32

Acknowledgements: We would like to thank Andrew Harvey, Roderick McCrorie, Peter Phillips, Enrique Sentana, two anonymous referees, and participants of the Recent Developments in Financial Econometrics and Empirical Finance Conference hosted by the Essex Centre for Financial Econometrics at the University of Essex, June 12–13, 2014, as well as seminar participants at the University of Reading, for helpful comments.

Address for Correspondence: Michael A. Thornton, Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, England.
Tel: +44 1904 324566; fax: +44 1904 ; e-mail:michael.thornton@york.ac.uk.

1. Introduction

Much of the theoretical framework of modern finance is expressed in continuous time. Since the seminal work by Merton (1969) on optimal portfolio choice under uncertainty, continuous time models have been used to study a wide range of applications including, among many others: the pricing of derivatives (Black and Scholes, 1973; Merton, 1973); the term structure of interest rates (Vasicek, 1977; Brennan and Schwartz, 1979); and asset pricing (Huang, 1987). Recent comprehensive reviews of the field can be found, for example, in Andersen, Davis, Kreiss and Mikosch (2009) and Aït-Sahalia and Hansen (2010).[1] Naturally, interest in methods for estimating continuous time models in finance has grown alongside developments in theory. The financial econometrician is, however, almost always forced to work with discretely observed data, continuous time financial data being either unavailable or contaminated with excessive micro-structure noise, and so estimation typically rests upon the transformation of the theoretical model into some analogous discrete time form.

To date, much of the estimation of continuous time models in finance has been restricted to Markov processes, which hold out the promise of estimation by maximum likelihood provided a series of transition probability densities can be estimated. Unfortunately, for many Markov systems there is no closed form solution for these densities and maximum likelihood estimation must then be based on an approximation technique; see the discussion in Phillips and Yu (2009). The simplest is the Euler approximation, in which (unobservable) derivatives with respect to time are replaced with (observable) differences over time. While it has been shown by Bergstrom (1984) for linear diffusions and by Florens-Zmirou (1989) for more general diffusions that estimators based on this approximation converge to the true values as the sampling interval shrinks to zero, it has been shown by Lo (1988), in the univariate case, and Wang, Phillips and Yu (2011), for multivariate diffusions, that estimates are inconsistent when the sampling interval remains constant. Therefore, unless data are observed at a sufficiently high frequency, typically daily or better, a more accurate approximation is needed.[2] However, the use of high frequency data comes not entirely without cost, because it is likely that problems of microstructure noise and/or unequally-spaced data must also be confronted.

In this paper we turn our attention to the estimation of the class of linear continuous time models, which include Markov processes as a special case. This class of continuous time autoregressive moving average (CARMA) models has been discussed by Brockwell (2001) and Brockwell and Marquardt (2005) and has been applied to interest rates by Andresen, Benth, Koekebakker and Zakamulin (2014) and to electricity futures by Benth and Šaltytė Benth (2009), while Brockwell (2004, 2009) discusses their application to GARCH and stochastic volatility models. We explore three benchmark models in finance using recently developed techniques of Chambers and Thornton (2012, 2016) to formulate the exact discrete representation of the CARMA model, that is to say the discrete time process which matches perfectly the first and second moments of the discretely observed process generated by the CARMA model and which lends itself to estimation using the Gaussian quasi-maximum likelihood method proposed by Bergstrom (1983). The advantages of this method, including

[1] Both of these collections of articles also cover material that goes beyond the specific focus of the material covered here.

[2] Methods based on Hermite polynomials have been proposed by Aït-Sahalia (2002, 2008), but these require numerical optimisation techniques.

computational efficiency, have been elegantly laid out in Bergstrom (1990) and in Bergstrom and Nowman (2007).

Alternative approaches to the estimation of linear continuous time systems have been proposed by Harvey and Stock (1985, 1988) and Zadrozny (1988), based on state space methods, and Robinson (1976, 1993), using spectral techniques. The state space approach employs Kalman filtering techniques to compute the Gaussian likelihood function. It avoids the derivation of the exact discrete time representation by producing optimal filtered estimates of the unobservable components in the state space form. But, as argued in Bergstrom (1990), the filtering approach imposes a higher computational burden than does the method based on the exact representation, and it is of interest, and can be important, to have knowledge concerning the dynamic evolution of the discrete time observations. For example, it is not clear from the state space approach that point-in-time observations generated by a CARMA(p, q) system satisfy an ARMA$(p, p-1)$ representation in discrete time, a feature that is clear from the exact model; see, for example, Chambers and Thornton (2012). However, the state space approach is particularly amenable to handling data irregularities, such as irregular sampling intervals and data observed at mixed frequencies, features that lead to even greater complexity in deriving exact representation; see, for example, Chambers (2015) in the case of mixed frequency data. Frequency domain methods also avoid the need for deriving the exact discrete model and estimates of the model parameters are typically obtained by maximising the Whittle approximation to the Gaussian likelihood. Such methods have, however, found relatively few applications in (financial) econometrics.

The first application of CARMA models concerns short-term interest rates, generalising the univariate model of Vasicek (1977), while the second application is to a term structure model of the relationship between long and short rates, generalising the bivariate Brennan-Schwartz (1979) model. This research follows the use of the exact discrete representation of a continuous time AR(1) process by Nowman (1997, 1998) to estimate a range of models of the term structure for interest rates. We depart from the parametrisations used by Nowman by estimating models capable of displaying a more sophisticated covariance structure and in the construction of the volatility component, which here is taken to be constant. In both of these applications we use UK data and explore the robustness of the estimated parameters across different sampling frequencies, these being weekly, monthly and quarterly. We avoid using higher-frequency data so that we can abstract from additional complications such as microstructure noise, day-of-the-week effects etc. We also acknowledge that other models (e.g. non-linear ones) and methods are likely to perform better with high frequency data and refer the reader to Aït-Sahalia and Jacod (2014) for a comprehensive modern treatment. In both of these applications we find that higher order terms, including the moving average error, improve the fit of the model significantly.

In the third application we consider a cointegrated present value model of stock prices and dividends due to Campbell and Shiller (1987), translated into continuous time. We imagine there to be sufficient friction in the market to prevent arbitrage. Here we face an additional complication in the way the data are observed. Whereas stock price data are observed at points in time, data on dividends are (in effect) time aggregates of activity over the observation period. That is to say, we are modelling a situation in which, following Campbell and Kyle (1993), the market is pricing the stock continuously through time based on the observations it makes on firms' accrual of the ability to pay dividends; the paid dividend reflects the aggregate of that accrual. This issue is often overlooked in empirical finance[3]; see, for example, the discretisation in Sangvinatsos and Wachter (2005) and Koijen, Rodriguez and Sbuelz (2009). We also illustrate the importance of the (correct flow) treatment of dividends by re-estimating the model under the incorrect assumption that the data are pure stocks and comparing persistence in the residuals and the discrete analogues across the two treatments.

In all three applications our preferred continuous time statistical model is not Markov. It could be argued that, in certain cases, the models correspond to an underlying theoretical model that was Markov but non-linear. Many linear statistical models in finance result from the approximation of a non-linear function, such as in Brennan and Schwartz (1979) whose diffusion process results from log-linearisation around a steady state, or in the discussion of the present value model in Campbell and Shiller (1987) or Campbell, Lo and MacKinlay (1997). In such cases, higher order terms enable a better approximation to the underlying theoretical model, aiding forecasting.

At the same time, linear continuous time models with moving average disturbances may arise from the aggregation of independent linear processes in much the same way as they do in discrete time modelling. Our non-Markov CARMA models could be the reduced form representation of the affine combinations of linear Markov processes, with the econometrician denied access to the disaggregated factors. An example of this construction is given in section 3.3 and the Appendix. Many other (perhaps less formal) models in finance, however, such as the discussion of momentum and mean reversion in stock returns following the insights of Jegadeesh and Titman (1993), are deliberately not Markov.

This paper is organised as follows. Section 2 is concerned with the specification of CARMA systems and draws on the results of Chambers and Thornton (2012) and Thornton and Chambers (2016) concerning exact discrete time representations. It also summarises the estimation and testing methods employed in the empirical work that follows in the subsequent section, which contains our three applications defined above. Section 4 provides some concluding comments.

2. Continuous time ARMA models

Many continuous time models employed in the field of finance are specified as diffusions. For a scalar variable $x(t)$ a general form of parametric diffusion takes the form

$$dx(t) = \mu(x(t);\theta)dt + \sigma(x(t);\theta)dW(t), \quad t > 0, \tag{1}$$

where $\mu(x;\theta)$ and $\sigma(x;\theta)$ are known functions that depend on an unknown parameter vector θ, $W(t)$ is a Wiener process and $x(0)$ can be taken to be fixed. The function $\mu(x;\theta)$ is often referred to as the drift function while $\sigma(x;\theta)$ is known as the volatility or diffusion function. The variable $x(t)$ generated by (1) satisfies the stochastic integral equation

$$x(t) = x(0) + \int_0^t \mu(x(s);\theta)ds + \int_0^t \sigma(x(s);\theta)dW(s), \quad t > 0,$$

[3] We are grateful to Enrique Sentana for drawing our attention to this feature of the literature.

which can be used as a basis for the development of methods for estimating θ using a sample of observations at discrete points of time given by $x_h, x_{2h}, \ldots, x_{Th}$, where h denotes the sampling interval. A review of such methods, as well as nonparametric approaches where the drift and diffusion functions are of the form $\mu(x)$ and $\sigma(x)$, respectively, and are assumed unknown, can be found in Aït-Sahalia (2007).

In many applications the estimates of the diffusion process (1) are used for some subsequent task, such as the pricing of options or the extraction of volatility estimates. Relatively little attention appears to be paid, however, to questions of how well the model fits the data, which is something that is often done in many econometric applications. For example, it is common to carry out various (mis-)specification tests, an obvious one in the context of diffusions being how well the model captures the dynamic evolution of the variable of interest. Lagrange multiplier (LM) and portmanteau-type tests are widely used to detect the presence of serial correlation in the residuals of an estimated model, which can be an indicator of dynamic misspecification; this does not appear to be common practice in the estimation of diffusions, a notable exception being de los Rios and Sentana (2011).

In some empirical illustrations of CARMA models (including an application to a short-term interest rate) Chambers and Thornton (2012) found that higher-order ARMA dynamics in the continuous time model could dramatically improve the ability of the model to capture the dynamics present in the observed discrete time data. It is of some considerable interest to explore this finding more widely in the context of additional applications using financial data. In order to do so we first outline the specification of CARMA models in general before moving on to consider issues relating to estimation and testing. Specific applications of this methodology then follow in section 3.

2.1. Specification

The continuous time ARMA(p, q) model for the $n \times 1$ vector $x(t)$ is given by

$$D^p x(t) = a_0 + A_{p-1} D^{p-1} x(t) + \ldots + A_0 x(t) + u(t) + \Theta_1 D u(t) + \ldots + \Theta_q D^q u(t), \; t > 0, \quad (2)$$

where D denotes the mean square differential operator satisfying

$$\lim_{\delta \to 0} E \left\{ \frac{x(t+\delta) - x(t)}{\delta} - Dx(t) \right\}^2 = 0,$$

$A_0, \ldots, A_{p-1}$ and $\Theta_1, \ldots, \Theta_q$ are $n \times n$ matrices of unknown coefficients, a_0 is an $n \times 1$ vector of unknown constants, and $u(t)$ is an $n \times 1$ continuous time white noise vector with variance matrix Σ. The matrices of unknown coefficients may, of course, depend on an underlying parameter vector θ of more deeply embedded structural parameters, provided that the elements of the matrices are known functions of θ. Although the process $u(t)$ and its derivatives are not physically realizable, systems such as (2) are nevertheless of widespread interest, and the condition $q < p$ is imposed so that $x(t)$ itself has an integrable spectral density matrix and, hence, has finite variance. The task is to estimate the matrices $A_0, \ldots, A_{p-1}$ and $\Theta_1, \ldots, \Theta_q$ and the vector a_0 of unknown (finite) coefficients, plus the variance matrix Σ of the continuous time white noise vector $u(t)$, from a sample of discrete time observations.

In the most general case the data satisfying (2) will contain both stocks and flows.

Without loss of generality, we can partition the vector of interest as

$$x(t) = \begin{bmatrix} x^s(t) \\ x^f(t) \end{bmatrix},$$

where $x^s(t)$ $(n^s \times 1)$ contains stock variables, $x^f(t)$ $(n^f \times 1)$ contains flow variables, and $n^s + n^f = n$. These variables are observed in different ways. Stock variables (such as asset prices, interest rates, exchange rates) are observed at points in time, so that the observations are of the form $x^s_{th} = x^s(th)$ $(t = 1, \ldots, T)$, while flow variables (such as dividends, income, profits) are observed as accumulations of the underlying rate of flow during the observation interval, yielding

$$x^f_{th} = \int_{th-h}^{th} x^f(r)dr, \quad t = 1, \ldots, T.$$

The key to deriving an exact discrete time representation for the observations lies in manipulating the (mean square) solution to (2), a process which eliminates all the unobservable components (e.g. derivatives of x) and delivers a random disturbance whose correlation properties can be derived. Let x_{th} denote the observed vector. Chambers and Thornton (2012) and Thornton and Chambers (2016) show that the observations satisfy the system

$$x_{th} = f_0 + F_1 x_{th-h} + \ldots + F_p x_{th-ph} + \eta_{th}, \quad t = p+1, \ldots, T, \tag{3}$$

where the vector f and matrices $F_1, \ldots, F_p$ are functions of the autoregressive parameters of the continuous time system (2), and the autocovariances of η_{th} depend on both the autoregressive and moving average parameters of (2), including the variance matrix Σ. In fact, η_{th} is an MA$(p-1)$ process if $x_{th} = x^s_{th}$ i.e. comprises solely stock variables, and is MA(p) when $x_{th} = x^f_{th}$ or $x_{th} = [x^{s\prime}_{th}, x^{f\prime}_{th}]'$ i.e. in the case of pure flows or a mixture of stocks and flows. Not only does the exact discrete time representation (3) form a basis for estimating the parameters of the continuous time model but it can also be used for forecasting; see Bergstrom (1990, chapter 8) and Chambers (1991) for details of forecasting with exact discrete models.

2.2. Estimation and testing

The discrete time ARMA(p, $p-1$) or ARMA(p, p) representation in (3) that corresponds to the continuous time ARMA(p, q) system (2) forms a natural basis for estimation of the unknown parameters. It is convenient to let β denote the vector of unknown parameters which is comprised of the elements of a_0, $A_0, \ldots, A_{p-1}$, $\Theta_1, \ldots, \Theta_q$ and Σ. The Gaussian likelihood methods detailed in Bergstrom (1990) for CARMA(2,0) systems can naturally be extended to CARMA(p, q) systems as in Chambers and Thornton (2012) and Thornton and Chambers (2016). Let $\eta = [\eta'_{ph+h}, \eta'_{ph+2h}, \ldots, \eta'_{Th}]'$ denote the $nT^* \times 1$ vector of disturbances, where $T^* = T - p$ denotes the effective sample size once allowance has been made for the p lags in (3). The covariance matrix of η, $E(\eta\eta') = \Omega_\eta$, has a sparse Toeplitz structure (reflecting the MA form) whose elements are functions of both the continuous time MA and AR parameters as well as Σ. As this matrix is positive definite and symmetric we can find a lower triangular matrix, M, with i, j'th element m_{ij}, such that

$$MM' = \Omega_\eta.$$

Bergstrom (1990, chapter 7) showed that M also reflects the sparse nature of Ω_η and, moreover, its elements converge rapidly to constants as one moves deeper into the matrix, leading to considerable computational advantages. A recursive procedure can be used to produce a normalised vector e, satisfying $E(e) = 0$ and $E(ee') = I_{nT^*}$, such that $Me = \eta$. The Gaussian log-likelihood function can then be evaluated as

$$\log L(\beta) = -\frac{nT^*}{2}\log 2\pi - \frac{1}{2}\sum_{i=p+1}^{nT}\left(e_i^2 + 2\log m_{ii}\right),$$

and the Gaussian (quasi maximum likelihood) estimator, $\hat{\beta}$, is the argument that maximises $\log L(\beta)$. Under standard regularity conditions of the type outlined in Bergstrom (1983), the estimator $\hat{\beta}$ is consistent and asymptotically normally distributed, converging at the rate $T^{1/2}$ to the limit distribution.

The Gaussian log-likelihood provides a convenient vehicle for the testing of hypotheses about the parameter vector β. If $\hat{\beta}_r$ denotes the estimator of β subject to a set of (possibly nonlinear) restrictions then the likelihood ratio statistic

$$LR = -2\left[\log L(\hat{\beta}_r) - \log L(\hat{\beta})\right]$$

can, under appropriate regularity conditions, be expected to have an asymptotic χ^2_g distribution under the null hypothesis, where g denotes the number of restrictions being tested. In addition, the vector e used to compute the log-likelihood function can be used to conduct a general test of dynamic specification. Bergstrom (1990, chapter 7) proposed a portmanteau-type test statistic based on the vectors of normalised residuals, e_{th}; it is of the form

$$S_l = \frac{1}{n(T^* - l)}\sum_{r=1}^{l}\left(\sum_{t=l+1}^{T} e'_{th}e_{th-rh}\right)^2,$$

which, under the null hypothesis that the model is correctly specified, has an approximate χ^2_l distribution for sufficiently large l and $T^* - l$, where l ($> p$) denotes the number of lags used. As is common with portmanteau tests, a significant value of S_l suggests dynamic misspecification of some form, although it does not indicate the precise nature of the misspecification. In the context of CARMA(p, q) models it would typically suggest that either p or q or both were insufficiently large enough to capture the dynamics of the observed variable. In the empirical work reported below, both LR (applied to hypotheses of interest) and S_l are used to test the specification of the estimated models.[4]

3. Applications in finance

This section considers three applications of CARMA models to topics of interest in finance, namely a model of short-term interest rates, a model of the term structure of interest rates, and a present value model of stock prices and dividends. One of our principal aims is to examine the robustness of the estimated continuous time model parameters when

[4]The properties of these and other misspecification tests are the subject of ongoing work by the authors; see Chambers and Thornton (2016).

the discrete time observation frequency is allowed to vary. In all the tables of results the numbers reported in parentheses are standard errors while the entries for the statistics S_l are the relevant p-values i.e. the proportion of the χ_l^2 distribution lying to the right of S_l.

3.1. Short-term interest rates

Recent work, e.g. Andresen, Benth, Koekebakker and Zakamulin (2014), has suggested that CARMA models may be suitable representations for short-term interest rates. These authors propose a number of reasons why CARMA models may be preferable to the more commonly used first-order Vasicek-type models, not least the fact that they can provide a better empirical fit to the observed term structure dynamics. Defining $r(t)$ to be the interest rate under consideration, we shall consider the CARMA$(2, 1)$ model given by

$$D^2 r(t) = a_0 + A_1 Dr(t) + A_0 r(t) + u(t) + \theta Du(t), \; t > 0, \tag{4}$$

where a_0, A_1, A_0 and θ are scalar parameters, and $u(t)$ is a mean zero uncorrelated process with variance σ_u^2. In addition to the CARMA$(2, 1)$ model we also consider the CARMA$(2, 0)$ specification (obtained by setting $\theta = 0$) and the CARMA$(1, 0)$ model, given by

$$Dr(t) = a_0 + A_0 r(t) + u(t);$$

note that this model is not nested within (4) i.e. it is not possible to impose restrictions on the parameters of (4) to obtain the CARMA$(1, 0)$ above.

Daily data on the Sterling one-month mean interbank lending rate were obtained from the Bank of England for the period 3 January 1978 to 6 November 2008; the properties of the data show a significant change after this point due to the burgeoning financial crisis. The daily data were aggregated to weekly, monthly and quarterly sampling intervals with the aggregated observations being the appropriate end-of-period values so that the series are genuinely of the stock variety. The sampling interval, h, was normalised to unity for the quarterly frequency so that, for monthly data, $h = 1/3$ while for the weekly data, $h = 150/1985 = 0.0756$ (this is the number of quarterly observations divided by the number of weekly observations, which is close to $1/13 = 0.0769$).

Results for all three models for the three sampling frequencies are given in Table 1. The parameter estimates are relatively stable across sampling frequencies for the CARMA$(1, 0)$ and CARMA$(2, 1)$ models but less so for the CARMA$(2, 0)$ model. Likelihood ratio tests convincingly reject the null that $\theta = 0$ at the weekly and monthly sampling frequencies but not at the quarterly frequency. None of the portmanteau statistics is significant at the 5% level for any model at any sampling frequency although the p-values are largest for the CARMA$(2, 1)$ model.

The estimates reported in Table 1 are based on the exact discrete time model corresponding to the underlying continuous time process. It is of interest to compare these estimates with those obtained using an approximation method, and for this comparison we have chosen the Euler approximation method. This has three main components: (i) replacing time derivatives of $r(t)$ with a discrete approximation based on the observations, r_{th}; (ii) evaluating $r(t)$ in the differential equation at r_{th-h}; and (iii) treating $u(t)$ as $DW(t)$, where $W(t)$ is a Wiener process, and approximating this derivative in discrete time using an *iid* Normal variate. Precise details of the Euler approximation schemes as applied to the three CARMA models are provided in the Appendix and Table 2 contains the parameter estimates obtained using the Euler approximation scheme.

As can be seen from Table 2 the estimates in the case of the CARMA$(1,0)$ models are virtually indistinguishable from those using the exact discrete time representation. This is explained mainly by the fact that the autoregressive coefficients in the exact and approximate models are $e^{A_0 h}$ and $1 + A_0 h$, respectively, and that $e^{A_0 h} = 1 + A_0 h + O(A_0^2 h^2)$. For example, in the case of quarterly data, the exact method gives $e^{-0.0351} = 0.9655$ while the approximation results in $1 - 0.0345 = 0.9655$. The standard errors, however, for a_0 and A_0 are larger when using the Euler approximation. The estimates obtained using the approximate method show greater divergence from those based on the exact method for the CARMA$(2,0)$ and CARMA$(2,1)$ models, in which the relationships between the continuous time and discrete time coefficients are much more complicated. They also tend to display less robustness across frequencies than the estimates obtained using the exact discrete time representation for these second-order autoregressive specifications.

Another comparison between the two approaches can be made in terms of the estimated roots to the continuous time autoregressive polynomials. For the CARMA$(1,0)$ model the roots are of the equation $z - A_0 = 0$ while for the CARMA$(2,0)$ and CARMA$(2,1)$ models the equation of interest is $z^2 - A_1 z - A_0 = 0$; in both cases, the roots are required to have negative real parts for the estimated equation to be dynamically stable. The estimated roots are reported in Table 3. While all the roots satisfy the stability condition there are some notable differences between the two approaches, most notable for the CARMA$(2,0)$ model. Also, the roots for the CARMA$(2,1)$ model are a complex conjugate pair under the exact discrete time representation for all three sampling frequencies whereas under the Euler approximation only the roots with monthly data are complex.

One of the motivations for estimating models of short-term interest rates is for the purposes of pricing bonds (and other derivative securities) and deriving estimates of the yield curve. Suppose the (short-term) interest rate, $r(t)$, satisfies the CARMA(p,q) model

$$D^p r(t) = a_0 + A_{p-1} D^{p-1} r(t) + \ldots + A_1 D r(t) + A_0 r(t) + u(t) + \Theta_1 D u(t) + \ldots + \Theta_q D^q u(t),$$

where $u(t)$ can be regarded (heuristically) as having the same properties as $\sigma_u DW(t)$, where $W(t)$ is a Wiener process and σ_u is a positive scalar parameter. At time th, where h denotes the sampling interval, the price of a zero coupon bond paying one unit upon maturity at time $Th > th$ is given by

$$P(th, Th) = E\left[\exp\left(-\int_{th}^{Th} r(s)ds\right)\middle| I(th)\right] = \exp\left(-\mu(th, Th) + \frac{1}{2}\sigma^2(th, Th)\right),$$

where $I(th)$ denotes the information set at time th and $\mu(th, Th)$ and $\sigma^2(th, Th)$ denote the mean and variance, respectively, of $\int_{th}^{Th} r(s)ds$ under the risk-free measure. The *yield to maturity* from buying the bond at th and selling at Th is given by

$$\gamma(th, Th) = -\frac{\log P(th, Th)}{Th - th};$$

the short rate is simply $\lim_{T\downarrow t}\gamma(th, Th)$. Under risk-neutral pricing (or the local expectations hypothesis, under which the risk-neutral and data generating measures coincide)

$$\gamma(th, Th) = \frac{\mu(th, Th) - \frac{1}{2}\sigma^2(th, Th)}{Th - th}.$$

Precise expressions for $\mu(th, Th)$ and $\sigma^2(th, Th)$ in terms of the parameters of the CARMA(p, q) model are provided in the Appendix.

In order to assess how well CARMA models can fit empirical yields we follow the approach of Benth, Koekebakker and Zakamulin (2008) and use data published by the Bank of England on the estimated yield curve. Three different dates are chosen which correspond to different shapes of the yield curve, the dates being 31 March 1998, 31 January 2000 and 30 June 2007. Let $\Gamma(th, Th)$ denote the empirical yield and $\gamma(th, Th, \theta)$ denote the estimated yield using a CARMA model with parameter vector θ. We then choose $\tilde{\theta}$ as the solution to

$$\tilde{\theta} = \arg\min_{\theta} S(\theta) \quad \text{where} \quad S(\theta) = \sum_{t=1}^{T} \left(\Gamma(th, Th) - \gamma(th, Th, \theta)\right)^2,$$

given the observed short rate at $t = 0$. The results of this exercise are depicted in Figures 1–3 using a horizon of $T = 60$ months for CARMA$(1, 0)$ and CARMA$(2, 1)$ specifications. Both CARMA models provide a good representation of the empirical yields although the fit of the CARMA $(2, 1)$ is better than that of the CARMA$(1, 0)$ in two of the three cases – the minimised values of the objective functions are 0.0050 and 0.0669, respectively, in Figure 1, 0.0025 and 0.0095 in Figure 2, and 0.0018 and 0.0005 in Figure 3. This is, perhaps, not too unexpected in view of the CARMA$(2, 1)$ model having more parameters than the CARMA$(1, 0)$ with which to capture the shape of the relevant curve.

3.2 A bivariate term structure model

The log-linearised version of the Brennan and Schwartz (1979) model of the term structure of interest rates consists[5] of the following pair of stochastic differential equations for the short rate, $r(t)$, and the long rate, $l(t)$:

$$d\ln r(t) = \alpha\left[\ln l(t) - \ln r(t) - \ln p\right] dt + \sigma_1 dz_1(t), \tag{5}$$
$$d\ln l(t) = \left[q - k_1 \ln r(t) - k_2 \ln l(t)\right] dt + \sigma_2 dz_2(t), \tag{6}$$

where α, k_1, k_2, q, $\ln p$, σ_1 and σ_2 are unknown parameters, and $z_1(t)$ and $z_2(t)$ are Wiener processes (or standard Brownian motions) with unknown correlation parameter ρ.[6] In this model α is a speed-of-adjustment parameter, p is a target value for the ratio l/r, and σ_1^2 and σ_2^2 represent the variances of the shocks to the system. Defining $x(t) = [\ln r(t), \ln l(t)]'$ the system can be equivalently written in the form of the CARMA$(1, 0)$ model

$$Dx(t) = a_0 + A_0 x(t) + u(t), \tag{7}$$

[5]See equations (17) and (18) of Brennan and Schwartz (1979).

[6]Hence $E[dz_1(t)dz_2(t)] = \sigma_1\sigma_2\rho dt$.

where

$$a_0 = \begin{pmatrix} -\alpha \ln p \\ q \end{pmatrix}, \quad A_0 = \begin{pmatrix} -\alpha & \alpha \\ -k_1 & -k_2 \end{pmatrix}$$

and $u(t) \sim N(0, \Sigma_u)$ where

$$\Sigma_u = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

This is the form of system that has been widely estimated although the emphasis is usually on obtaining parameter estimates for use in the pricing of bonds (and options) rather than assessing how well the model captures the salient features of the data.

The main features of the model outlined above are easily migrated across to a more general CARMA specification. For example, a CARMA$(2,1)$ system can be specified as

$$D^2 x(t) = a_0 + A_1 Dx(t) + A_0 x(t) + u(t) + \Theta_1 Du(t), \tag{8}$$

in which a_0, A_0 and $u(t)$ are as defined above and where

$$A_1 = \begin{pmatrix} -\gamma_1 & -\gamma_3 \\ -\gamma_4 & -\gamma_2 \end{pmatrix}, \quad \Theta_1 = \begin{pmatrix} \theta_1 & \theta_3 \\ \theta_4 & \theta_2 \end{pmatrix}.$$

Setting the elements of Θ_1 to zero yields the CARMA$(2,0)$ system; such restrictions are easily tested using observed data. The two equations encapsulated in (8) are

$$\begin{aligned} D^2 \ln r(t) &= -\gamma_1 D \ln r(t) - \gamma_3 D \ln l(t) + \alpha \left[\ln l(t) - \ln r(t) - \ln p\right] + v_1(t), \\ D^2 \ln l(t) &= -\gamma_4 D \ln r(t) - \gamma_2 D \ln l(t) + \left[q - k_1 \ln r(t) - k_2 \ln l(t)\right] + v_2(t), \end{aligned}$$

where $v_1(t) = u_1(t) + \theta_1 Du_1(t) + \theta_3 Du_2(t)$ and $v_2(t) = u_2(t) + \theta_4 Du_1(t) + \theta_2 Du_2(t)$ for notational covenience. Interest rate equations in the form of second-order stochastic differential equations are not without precedent. A CARMA$(2,0)$ specification (in effect) was used in the continuous time macroeconometric model of the United Kingdom by Bergstrom and Nowman (2007) while Andresen, Benth, Koekebakker and Zakamulin (2014) have more recently developed more general CARMA specifications.

In the empirical work we take the short rate to be the Sterling one-month mean interbank lending rate and the long rate to be the yield on twenty year British Government securities with a nominal zero coupon. Daily data were obtained from the Bank of England for the period 11 February 1992 to 6 November 2008 and aggregated to weekly, monthly and quarterly frequencies.[7] Estimates of three continuous time models are given in Tables 4–6, these being CARMA$(1,0)$, CARMA$(2,0)$ and CARMA$(2,1)$ using weekly, monthly and quarterly data. In the second-order models we have imposed the constraints that $\gamma_3 = \gamma_4 = 0$ in the matrix A_1 and that $\theta_3 = \theta_4 = 0$ in the matrix of continuous time moving average coefficients, Θ_1. Also, for the CARMA$(2,0)$ and CARMA$(2,1)$ models we additionally set $q = \ln p = 0$.[8]

[7]The aggregated observations are the end-of-period values so that the series are genuinely of the stock variety.

[8]Some convergence problems were encountered without imposing this restriction using weekly data. The

Four particular aspects of the results are worth commenting on. First, all of the roots of the CARMA$(1, 0)$ models are real and positive, indicating that these estimated systems are unstable. In contrast all of the roots of the CARMA$(2, 0)$ and CARMA$(2, 1)$ systems have negative real parts and, hence, they are stable. Secondly, a certain amount of instability in the parameter estimates across frequencies can be detected, with some even changing signs. This could be interpreted as evidence against the validity of the underlying continuous time model(s). Thirdly, it can be seen that moving from the CARMA$(1, 0)$ to the CARMA$(2, 0)$ specification does not eradicate the evidence of dynamic misspecification at the weekly frequency and only marginally does so at the monthly frequency, as indicated by the p-values of the S_{12} statistics. This contrasts with the CARMA$(2, 1)$ model for which the the S_{12} statistics have large p-values at all frequencies. Finally, Table 6 reports the p-values for the likelihood ratio test of the null hypothesis that $\theta_1 = \theta_2 = 0$ i.e. of the CARMA$(2, 0)$ nested within the CARMA$(2, 1)$. The null hypothesis is clearly rejected at the 5% level of significance at all sampling frequencies further supporting the conjecture that the presence of the MA component in the continuous time system has empirical content.

3.3. A present value model of stock prices and dividends

Present value models stipulate that, in the absence of long-run bubbles, stock prices should represent the discounted flow of future dividends and that this leads to a long run relationship between real stock prices and real dividends. Campbell and Shiller (1987, p.17) note the differences in timing between the two series, with stock prices measured 'beginning-of-period' and a dividend 'paid some time within period t.' In their discrete time model, Campbell and Shiller were concerned about what might be known about the current period's dividend payout when the stock price was measured, and they constructed composite variables as controls. The advantage of working in continuous time is that revelations in information within observation periods are modelled explicitly. Much of the literature since, however (see, for example, Sangvinatsos and Wachter, 2005, and Koijen, Rodriguez and Sbuelz, 2009), has tended to ignore this distinction and to regard data on the stock price, $s(t)$, and on dividends, $d(t)$, as consisting of the value of a continuous time process at a specific point in time. This is as if firms were paying dividends continuously through time, but only those paid at a particular point in time were recorded, which is patently not the case, since there are no unobserved dividend payouts. In our baseline treatment we regard dividends as a flow, with the observed payout, $d_t = \int_{t-1}^{t} d(\tau)d\tau$, reflecting an observed accrual of profits over the observation period, while stock prices are modelled point-in-time, $s_t = s(t)$, both for $t = 1, 2, \ldots, T$. Estimated models using the conventional (but incorrect) treatment are also reported for comparison.[9]

Our approach mirrors that of Campbell and Kyle (1993), who estimate a continuous time model of stock prices and dividends, taking care to treat the dividends as a flow variable. After exponential de-trending they model[10] dividends, $d(t) = d_0(t) + d_1(t)$, as the sum of

null hypothesis was not rejected for the monthly and quarterly data and hence we report estimates for all frequencies with the restriction imposed. The restriction implies that the target value for l/r is equal to one.

[9]As we are only using data at a single sampling frequency in this application we set $h = 1$ throughout this sub-section.

[10]Model A, equation (2.5).

two independent linear processes,

$$\begin{aligned} Dd_0(t) &= \sigma_0 dz_0(t), \\ Dd_1(t) &= \alpha_1 d_1(t) + \sigma_1 dz_1(t), \end{aligned} \tag{9}$$

with $dz_0(t)$ and $dz_1(t)$ independent standard Brownian motions. It can be shown (see the Appendix) that $d(t)$ follows a continuous time autoregressive integrated moving average, or CARIMA$(1,1,1)$, model of the form

$$D^2 d(t) = \alpha_1 Dd(t) + u(t) + \theta_1 Du(t), \tag{10}$$

suggesting that an MA component may be considered as arising naturally in this framework. We build upon this feature in the following analysis.

In Campbell and Kyle (1993), only smart investors are able to discern the more persistent process $d_1(t)$ from the purely transitory process $d_0(t)$, making the decomposition of the observed dividends something of independent interest and naturally suitable for the application of the Kalman filter. The computational efficiency of our approach stems from avoiding such a decomposition, but a linear filter to perform one could be constructed from the estimated parameters. In a world of complete information and frictionless markets, prices following smooth continuous sample paths of bounded variation admit arbitrage opportunities; see Harrison, Pitbladdo and Schaefer (1984). While this remains an important benchmark, there is a growing body of research into models that violate this condition. Besides Campbell and Kyle (1993), arbitrage opportunities are present in the models of: Willard and Dybvig (1999), where the market constrains investors from making incredible promises in states that they believe will not occur; Basak and Croitoru (2000), where heterogeneity between agents can generate mispricing in equilibrium; Liu and Longstaff (2004), where risk aversion and the possibility of loss of collateral prevents investors from exploiting opportunities fully; Jarrow and Protter (2005), due to the presence of influential large traders; and, Koijen, Rodriguez and Sbuelz, (2009), who consider portfolio choice when stock prices exhibit momentum.

After taking logs, both series display unit-root type behaviour leading Campbell and Shiller (1987) to postulate that the long run relationship between the two series is a form of cointegration, with the discount factor determining the cointegrating vector. Following their work we analyse the relationship between the logarithm of the stock price and the logarithm of dividends using the same monthly data spanning the period 1871–1986,[11] which avoids the need to include share buy-backs as part of investor remuneration.

In the linear continuous time framework, $1 \leq r < n$ cointegrating relationships between the components of $x(t)$ imply that the $n \times n$ matrix A_0 has rank r and can be written in the form $A_0 = \alpha\beta'$, where α and β are both $n \times r$. The matrix α has the interpretation of containing speed-of-adjustment parameters while β is the matrix of cointegrating vectors such that $\beta' x(t)$ is stationary. As shown by Phillips (1991), the matrix of long-run parameters, β, is unaffected by the aliasing phenomenon; it contains cointegrating vectors of the observed data x_t, whether the variables are stocks or flows. The mixed stock-flow nature of the data means that, even for the simplest models, the short run parameters cannot be estimated by

[11] The data are available at http://www.econ.yale.edu/~shiller/data.htm.

a conventional VAR.

We consider three models based on equation (2) with $x(t) = (s(t), d(t))'$. In a two variable system, cointegration implies that we may write, without loss of generality, $A_0 = \alpha\beta'$, where $\alpha' = [\alpha_1, \alpha_2]$ and $\beta' = [1, \beta_1]$. In each case a value of β_1 in the vicinity of -1 is expected, with divergence the result of discounting of future dividends, while error correction implies $\alpha_1 < 0$ and $\alpha_2 > 0$. Estimates for the CARMA$(1, 0)$ model,

$$\begin{aligned} Ds(t) &= a_{0,1} + \alpha_1 s(t) + \alpha_1\beta_1 d(t) + u_1(t), \\ Dd(t) &= a_{0,2} + \alpha_2 s(t) + \alpha_2\beta_1 d(t) + u_2(t), \end{aligned}$$

where $u(t) = [u_1(t), u_2(t)]' \sim N(0, \Sigma_u)$ and $\Sigma_u = QQ'$ with Q a lower triangular matrix, are reported in Table 7. The estimate of β_1 is close to -1.4, but α_1 has the wrong sign, placing the burden of error correction within the system on dividends. The Bergstrom S statistic is in the extreme right tail of its asymptotic distribution for both 4 and 12 lags, suggesting a higher order dynamic structure is needed.

We also report, in Tables 8 and 9, estimates of CARMA$(2, 0)$ and CARMA$(2, 1)$ systems, respectively; the latter is given by

$$D^2 s(t) = a_{0,1} + A_{1,11} Ds(t) + A_{1,12} Dd(t) + \alpha_1 s(t) + \alpha_1\beta_1 d(t) + w_1(t), \tag{11}$$

$$D^2 d(t) = a_{0,2} + A_{1,21} Ds(t) + A_{1,22} Dd(t) + \alpha_2 s(t) + \alpha_2\beta_1 d(t) + w_2(t), \tag{12}$$

where $w_1(t) = u_1(t) + \Theta_{11} Du_1(t) + \Theta_{12} Du_2(t)$ and $w_2(t) = u_2(t) + \Theta_{21} Du_1(t) + \Theta_{22} Du_2(t)$ are defined for notational convenience. The CARMA$(2, 0)$ model is obtained by setting $\Theta_{i,j} = 0$ $(i, j = 1, 2)$. In both specifications the estimate of β_1 remains between -1.4 and -1.5 and the adjustment parameters in α have the expected sign. The evidence of dynamic misspecification given by the Bergstrom S statistic remains in the CARMA$(2, 0)$ model, with the CARMA$(2, 1)$ showing significant improvement over both purely autoregressive models. The CARMA$(2, 1)$ also has by far the highest log-likelihood and the test statistic for the likelihood ratio test of the restriction that the four continuous time MA parameters are jointly zero is over 680, far into the extreme tail of the asymptotic χ^2_4 distribution. The parameter Θ_{22} has by far the highest t-ratio, suggesting that it is the equation describing the law of motion for dividends that benefits most from the inclusion of a moving average error.

We now consider the effects of treating dividends, incorrectly, as a stock variable; that is to say, as if our dividend data were of the form $d_t = d(t)$ $(t = 1, 2, \ldots, T)$. Tables 10 to 12 report estimates for the above models under this treatment. Two features are worth reporting. First, the CARMA $(2, 1)$ model performs better than the two simpler versions in the plausibility that its errors are white noise and that the model where dividends are treated as a flow have the higher log-likelihood. Secondly, estimates of the cointegrating vector are remarkably consistent across both treatments. The effect of the different treatments is, not surprisingly, seen most clearly in the off-diagonal elements of A_1, which reflect the short-run impact of (time derivatives of) the two variables on one another, with their impact on themselves remaining relatively stable.

To compare the relative success of each model in explaining the dynamic relationship between stock prices and dividends, Figure 4 plots the autocorrelations and cross-correlations between the normalised residuals, e_t, generated by our six candidate models. For a correctly specified model, these should be independent white noise processes, and the autocorrelations explore the findings of the Bergstrom S statistic in greater detail. The first panel shows the autocorrelations for the normalised residuals on the stock equation, where both CARMA $(1,0)$ specifications exhibit relatively large first order autocorrelations of around 0.3, providing evidence of under-parametrisation, but there is little to chose between the other specifications. This is not true of the remaining panels. The second panel shows the cross-correlation between the normalised residual on the stock equation and lags of the normalised residual on the dividend equation. There is a noticeable up-tick in all series at the twelfth lag, suggesting that unusually high (low) stock prices might be related to unusually large (small) dividend payments one year previously in a way that these parsimonious models are not able to capture. That being said, the mixed CARMA$(2,1)$ model out-performs the others, including the corresponding pure stock model, particularly at short lags. Not surprisingly, the largest distinction between the treatments can be seen in the final panel, showing the autocorrelations for the normalised residuals on the dividend equation. It is clear that the normalised residuals produced by the CARMA$(1,0)$ models suffer from positive first order serial correlation, while the higher order pure stock models, including the CARMA $(2,1)$ exhibit negative first order serial correlation before bouncing back. Both CARMA$(2,0)$ models over-shoot zero for the second order autocorrelation. Only the normalised residuals from the mixed CARMA $(2,1)$, have the expected correlation structure. The normalised residuals from the pure stock CARMA $(2,1)$ do not, despite using the same number of parameters.

The implications of the two treatments of dividends are also revealed by the corresponding exact discrete time models. When data generated by equations (11) and (12) are observed at discrete intervals they have a vector error correction representation of the form

$$\begin{aligned}
\Delta s_t &= f_1 + a_1[s_{t-1} + \beta_1 d_{t-1}] + F_{11}\Delta s_{t-1} + F_{12}\Delta d_{t-1} + \eta_{1,t}, \\
\Delta d_t &= f_2 + a_2[s_{t-1} + \beta_1 d_{t-1}] + F_{21}\Delta s_{t-1} + F_{22}\Delta d_{t-1} + \eta_{2,t},
\end{aligned}$$

where Δ is the difference operator and $\eta_t = (\eta_{1,t}, \eta_{2,t})'$ has a moving average representation of order two when dividends are treated as a flow and order one when treated as a stock. Following Phillips (1991), the parameter β_1, describing the cointegrating relationship between the series, is the same as in (11) and (12). Table 13 presents the translation of the intercept and short-run parameters in the two treatments, where $a = [a_1, a_2]'$ denotes the vector of adjustment parameters. While signs remain unaltered, the magnitudes of some short-run parameters are clearly affected. The misspecified stock treatment of dividends noticeably underestimates the short-run impact of changes in the stock price on future dividends while overestimating the short-run impact of changes in dividends on future stock prices.

4. Concluding comments

This paper has considered the estimation of CARMA models in finance using their exact discrete time representation. CARMA systems offer a number of attractive features, including the ability to capture adequately the dynamics inherent in observed financial time series as well as being able to model stock and flow data in accordance with the different ways in which they are observed, while maintaining a closed-form expression for the likelihood. These features have been evident in our three empirical examples, and we have also shown that mistakenly treating a flow series as a stock has significant consequences for the estimates of parameters dictating short-run dynamics.

We have applied our techniques to relatively smooth time series that can be modelled as stochastic differential equations but a number of extensions are possible. The CARMA model driven by a Lévy process has already been discussed by Brockwell (2001) and Brockwell and Marquardt (2005) to model particularly volatile financial time series. Incorporating jump processes into our exact discrete framework would also extend the range of estimation techniques open to applied researchers. Such extensions are, however, beyond the scope of the current contribution.

Appendix

Derivation of the exact discrete time representation

We give a brief overview of the method for deriving the exact discrete representation of a continuous time ARMA process when the observed data, $x_{th} = x(th)$, are stocks. Extensions to processes involving flow data are covered in Chambers and Thornton (2012) and Thornton and Chambers (2016). We begin by noting that (2) can be written in state space form as

$$Dy(t) = a + Ay(t) + \Theta u(t), \tag{13}$$

where $y(t) = [y_1(t), y_2(t)', \ldots, y_p(t)']'$ with $x(t) = y_1(t)$ and

$$a = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_0 \end{pmatrix}, \quad A = \begin{pmatrix} A_{p-1} & I & 0 & \ldots & 0 \\ A_{p-2} & 0 & I & \ldots & 0 \\ \vdots & & & & \vdots \\ A_1 & 0 & 0 & \ldots & I \\ A_0 & 0 & 0 & \ldots & 0 \end{pmatrix}, \quad \Theta = \begin{pmatrix} \Theta_{p-1} \\ \Theta_{p-2} \\ \vdots \\ \Theta_1 \\ I \end{pmatrix},$$

where $\Theta_j = 0$ for $j > q$. Integrating (13) over $(0, th]$ gives

$$\begin{aligned} y(th) &= e^{Ath}y(0) + \int_0^{th} e^{A(th-s)} \left[a + \Theta u(s)\right] ds \\ &= e^{Aht}y(0) + \int_0^{t} e^{Ah(t-r)} \left[a + \Theta u(rh)\right] hdr, \quad t > 0, \end{aligned} \tag{14}$$

from which it follows that

$$y(th) = c + Cy((t-1)h) + \epsilon(th), \quad t = 1, \ldots, T, \tag{15}$$

where

$$c = \left(\int_0^1 C(r)dr\right) ah, \quad \epsilon(th) = \int_{(t-1)h}^{th} e^{A(th-s)}\Theta u(ds),$$

$C(r) = e^{rAh}$ and $C = e^{Ah}$. The disturbance vector in (15) has expectation zero and second moments given by

$$E\left(\epsilon(th)\epsilon(th)'\right) = \int_{(t-1)h}^{th} e^{A(th-s)}\Theta\Sigma\Theta' e^{A'(th-s)}ds = h\int_0^1 e^{Ahr}\Theta\Sigma\Theta' e^{A'hr}\, dr.$$

The exact discrete representation follows from deploying lags of (15) to solve for the unobservable elements in the state vector, $y(th)$, and eliminating them from the equation for $x(t)$ using the methods discussed in Chambers and Thornton (2012). Similar methods apply when the observations are on flow variables or a mixture of stocks and flows.

Derivation of the Euler approximation

The main steps in deriving the Euler approximation were outlined in the text. For the CARMA$(1,0)$ model,

$$Dr(t) = a_0 + A_0 r(t) + u(t),$$

the derivative on the left-hand side is approximated using

$$Dr(t) \approx \frac{r_{th} - r_{th-h}}{h}.$$

Noting that $DW(t) \sim N(0, \sigma_u^2/dt)$,[12] $u(t)$ can be approximated with

$$u(t) \approx \frac{\sigma_u}{h^{1/2}} e_{th} \sim N\left(0, \frac{\sigma_u^2}{h}\right),$$

where $e_{th} \sim iid\, N(0,1)$ and dt is approximated with h. The approximate model is then, replacing $r(t)$ with r_{th-h},

$$\frac{r_{th} - r_{th-h}}{h} = a_0 + A_0 r_{th-h} + \frac{\sigma_u}{h^{1/2}} e_{th},$$

which, upon rearranging, becomes

$$r_{th} = a_0 h + (1 + A_0 h) r_{th-h} + \sigma_u h^{1/2} e_{th},$$

so that r_{th} is ARMA(1,0) in discrete time.

Turning to the CARMA$(2,0)$ model

$$D^2 r(t) = a_0 + A_1 Dr(t) + A_0 r(t) + u(t)$$

we need the additional approximation

$$D^2 r(t) \approx \left(\frac{r_{th} - r_{th-h}}{h^2}\right) - \left(\frac{r_{th-h} - r_{th-2h}}{h^2}\right) = \frac{1}{h^2}\left(r_{th} - 2r_{th-h} + r_{th-2h}\right).$$

Again we evaluate terms involving $r(t)$ on the right-hand side at the preceding sample point to obtain

$$\frac{1}{h^2}\left(r_{th} - 2r_{th-h} + r_{th-2h}\right) = a_0 + A_1\left(\frac{r_{th-h} - r_{th-2h}}{h}\right) + A_0 r_{th-h} + \frac{\sigma_u}{h^{1/2}} e_{th}$$

[12] This is based on the observation that $dW(t) \sim N(0, \sigma_u^2 dt)$ and noting that $DW(t) = dW(t)/dt$, thereby introducing dt into the demoninator of the variance.

which can be rearranged to give the discrete time ARMA(2,0) representation

$$r_{th} = a_0 h^2 + (2 + A_1 h + A_0 h^2) r_{th-h} - (1 + A_1 h) r_{th-2h} + \sigma_u h^{3/2} e_{th}.$$

Finally, the CARMA$(2,1)$ model

$$D^2 r(t) = a_0 + A_1 Dr(t) + A_0 r(t) + u(t) + \theta Du(t)$$

requires the additional approximation for $Du(t)$, for which we use

$$Du(t) \approx \frac{\sigma_u}{h^{1/2}} \left(\frac{e_{th} - e_{th-h}}{h} \right) = \frac{\sigma_u}{h^{3/2}} \left(e_{th} - e_{th-h} \right).$$

Making the substitutions yields

$$\frac{1}{h^2} \left(r_{th} - 2r_{th-h} + r_{th-2h} \right) = a_0 + A_1 \left(\frac{r_{th-h} - r_{th-2h}}{h} \right) + A_0 r_{th-h} + \frac{\sigma_u}{h^{1/2}} e_{th} + \frac{\theta \sigma_u}{h^{3/2}} \left(e_{th} - e_{th-h} \right)$$

which can be rearranged to give

$$r_{th} = a_0 h^2 + (2 + A_1 h + A_0 h^2) r_{th-h} - (1 + A_1 h) r_{th-2h} + \sigma_u \left(h^{3/2} + \theta h^{1/2} \right) e_{th} - \sigma_u \theta h^{1/2} e_{th-h},$$

which is ARMA$(2,1)$ in discrete time.

Derivation of bond pricing formulae

In order to evaluate the components of $\gamma(th, Th)$ it is convenient to write the CARMA model in state space form. To do this we define the $p \times 1$ state vector $y(t) = [y_1(t), \ldots, y_p(t)]'$ and let $y_1(t) = r(t)$. It can then be shown that the CARMA(p, q) model for $r(t)$ satisfies the state space representation in (13). Let $S_1 = (1, 0, \ldots, 0)$ denote the $1 \times p$ selection vector that picks $r(t)$ from $y(t)$ i.e. $r(t) = S_1 y(t)$. Then the (conditional) mean and variance of $r(t)$ are given by the (conditional) mean and variance of $S_1 y(t)$.

We begin by noting that, for $s > th$,

$$y(s) = e^{A(s-th)} y(th) + \int_{th}^{s} e^{A(s-v)} \left(a + \Theta u(v) \right) dv$$

and so it follows that

$$\begin{aligned}
\int_{th}^{Th} y(s) ds &= \left(\int_{th}^{Th} e^{A(s-th)} ds \right) y(th) + \int_{th}^{Th} \int_{th}^{s} e^{A(s-v)} \left(a + \Theta u(v) \right) dv ds \\
&= \left(\int_{th}^{Th} e^{A(s-th)} ds \right) y(th) + \int_{th}^{Th} \left(\int_{th}^{s} e^{A(s-v)} dv \right) ds \cdot a \\
&\quad + \int_{th}^{Th} \int_{th}^{s} e^{A(s-v)} \Theta u(v) dv ds.
\end{aligned}$$

In what follows it is convenient to define the following integrals of the matrix exponential:

$$\Phi(x) = \int_0^x e^{As} ds, \quad \Upsilon(x) = \int_0^x \Phi(r) dr = \int_0^x \int_0^r e^{As} ds dr.$$

Taking the first component, a change of variable yields

$$\int_{th}^{Th} e^{A(s-th)} ds = \int_0^{Th-th} e^{Aw} dw = \Phi(Th - th).$$

Similarly, we also find that

$$\begin{aligned}
\int_{th}^{Th} \left(\int_{th}^{s} e^{A(s-v)} dv \right) ds &= \int_{th}^{Th} \left(\int_0^{s-th} e^{Aw} dw \right) ds \\
&= \int_{th}^{Th} \Phi(s - th) ds \\
&= \int_0^{Th-th} \Phi(x) dx = \Upsilon(Th - th).
\end{aligned}$$

For the stochastic integral involving $u(t)$ we can show that

$$\begin{aligned}
\int_{th}^{Th} \int_{th}^{s} e^{A(s-v)} \Theta u(v) dv ds &= \int_{th}^{Th} \int_{v}^{Th} e^{A(s-v)} \Theta u(v) ds dv \\
&= \int_{th}^{Th} \left(\int_{v}^{Th} e^{A(s-v)} ds \right) \Theta u(v) dv \\
&= \int_{th}^{Th} \left(\int_0^{Th-v} e^{Aw} dw \right) \Theta u(v) dv \\
&= \int_{th}^{Th} \Phi(Th - v) \Theta u(v) dv.
\end{aligned}$$

Hence we have shown that

$$\int_{th}^{Th} y(s) ds = \Phi(Th - th) y(th) + \Upsilon(Th - th) a + \int_{th}^{Th} \Phi(Th - v) \Theta u(v) dv.$$

From the usual rules of expectation and variance we obtain

$$\mu(th, Th) = S_1[\Phi(Th - th) y(th) + \Upsilon(Th - th) a],$$

$$\sigma^2(th, Th) = S_1 \int_{th}^{Th} \Phi(Th - v) \Sigma_\Theta \Phi(Th - v)' dv S_1',$$

where $\Sigma_\Theta = \sigma_u^2 \Theta \Theta'$. The integral in the last expression also simplifies using a change of variable:

$$\begin{aligned}
\int_{th}^{Th} \Phi(Th-v)\Sigma_\Theta \Phi(Th-v)' dv &= \int_0^{Th-th} \Phi(w)\Sigma_\Theta \Phi(w)' dw \\
&= \int_0^{Th-th} \left(\int_0^w e^{As} ds \right) \Sigma_\Theta \left(\int_0^w e^{Ar} dr \right)' dw \\
&= \int_0^{Th-th} \int_0^w \int_0^w e^{As} \Sigma_\Theta e^{A'r} dr ds dw \\
&= \Lambda(Th-th)
\end{aligned}$$

where the triple matrix exponential integral function $\Lambda(x)$ is defined by

$$\Lambda(x) = \int_0^x \int_0^w \int_0^w e^{As} \Sigma_\Theta e^{A'r} dr ds dw.$$

Using the above expressions it follows that

$$P(th, Th) = \exp\left(\frac{1}{2} S_1 \Lambda(Th-th) S_1' - S_1 \left[\Phi(Th-th) y(th) + \Upsilon(Th-th) a \right] \right)$$

and, hence, the yields are of the form

$$\gamma(th, Th) = \frac{1}{Th-th} \left(S_1 \left[\Phi(Th-th) y(th) + \Upsilon(Th-th) a \right] - \frac{1}{2} S_1 \Lambda(Th-th) S_1' \right).$$

Computationally, all of the integrals of the matrix exponential, and the matrix exponential itself, can be obtained from the computation of a single matrix exponential. Van Loan (1978) considered a matrix M, and its exponential $N(t) = e^{Mt}$, of the form

$$M = \begin{pmatrix} A_1 & B_1 & C_1 & D_1 \\ 0 & A_2 & B_2 & C_2 \\ 0 & 0 & A_3 & B_3 \\ 0 & 0 & 0 & A_4 \end{pmatrix}, \quad N(t) = \begin{pmatrix} F_1(t) & G_1(t) & H_1(t) & K_1(t) \\ 0 & F_2(t) & G_2(t) & H_2(t) \\ 0 & 0 & F_3(t) & G_3(t) \\ 0 & 0 & 0 & F_4(t) \end{pmatrix}.$$

By noting that $(d/dt)e^{Mt} = Me^{Mt}$ and solving subject to $e^{Mt}|_{t=0} = I$ he was able to show that

$$
\begin{aligned}
F_j(t) &= e^{A_j t}, \quad j = 1, \ldots, 4, \\
G_j(t) &= \int_0^t e^{A_j(t-s)} B_j e^{A_{j+1}s} ds, \quad j = 1, 2, 3, \\
H_j(t) &= \int_0^t e^{A_j(t-s)} C_j e^{A_{j+2}s} ds + \int_0^t \int_0^s e^{A_j(t-s)} B_j e^{A_{j+1}(s-r)} B_{j+1} e^{A_{j+2}r} drds, \quad j = 1, 2, \\
K_1(t) &= \int_0^t e^{A_1(t-s)} D_1 e^{A_4 s} ds + \int_0^t \int_0^s e^{A_1(t-s)} \left[C_1 e^{A_3(s-r)} B_3 + B_1 e^{A_2(s-r)} C_2 \right] e^{A_4 r} drds \\
&\quad + \int_0^t \int_0^s \int_0^r e^{A_1(t-s)} B_1 e^{A_2(s-r)} B_2 e^{A_3(r-w)} B_3 e^{A_4 w} dwdrds.
\end{aligned}
$$

By suitable choice of the sub-matrices of M it is possible to derive the functions of interest here from $e^{M(Th-th)}$. Specifically, taking

$$
M = \begin{pmatrix} -A & I & 0 & 0 \\ 0 & -A & I & 0 \\ 0 & 0 & 0 & \Sigma_\Theta \\ 0 & 0 & 0 & A' \end{pmatrix},
$$

we find, in particular, that

$$
\begin{aligned}
F_4(t) &= e^{A't}, \\
G_2(t) &= \int_0^t e^{-A(t-s)} ds = e^{-At} \int_0^t e^{As} ds, \\
H_1(t) &= \int_0^t \int_0^s e^{-A(t-s)} e^{-A(s-r)} drds = e^{-At} \int_0^t \int_0^s e^{Ar} drds, \\
K_1(t) &= \int_0^t \int_0^s \int_0^r e^{-A(t-s)} e^{-A(s-r)} \Sigma_\Theta e^{A'w} dwdrds \\
&= e^{-At} \int_0^t \int_0^s \int_0^r e^{Ar} \Sigma_\Theta e^{A'w} dwdrds.
\end{aligned}
$$

It then follows that

$$
\begin{aligned}
\Phi(t) &= F_4(t)' G_2(t), \\
\Upsilon(t) &= F_4(t)' H_1(t), \\
\Lambda(t) &= F_4(t)' K_1(t) + K_1(t)' F_4(t);
\end{aligned}
$$

see, also, Jewitt and McCrorie (2005, p.401) for a similar result. Note that $K_1(t)$ involves the triple integral with limits $\int_0^t \int_0^s \int_0^r \ldots dwdrds$ whereas $\Lambda(t)$ involves $\int_0^t \int_0^s \int_0^s \ldots dwdrds$ – the appearance of $K_1(t)$ and $K_1(t)'$ in the expression for $\Lambda(t)$ accounts for these differences.

Derivation of equation (10)

We express (9) as

$$\begin{aligned} Dd_0(t) &= \sigma_0 dz_0(t), \\ (D-\alpha_1)d_1(t) &= \sigma_1 dz_1(t). \end{aligned}$$

Multiplying the top equation through by $(D-\alpha_1)$ and the bottom by D before summing produces

$$(D-\alpha_1)Dd(t) = (D-\alpha_1)\sigma_0 dz_0(t) + D\sigma_1 dz_1(t).$$

The most straightforward way to demonstrate equivalence between the zero mean disturbance process $(D-\alpha_1)\sigma_0 dz_0(t) + D\sigma_1 dz_1(t)$ and the zero mean disturbance process on the right hand side of (10) is to match the spectra, see Priestley (1981), of the two processes,

$$\begin{aligned} (i\lambda-\alpha_1)(-i\lambda-\alpha_1)\frac{\sigma_0^2}{2\pi} + (i\lambda)(-i\lambda)\frac{\sigma_1^2}{2\pi} &= (1+\theta_1 i\lambda)(1-\theta_1 i\lambda)\frac{\sigma_u^2}{2\pi} \\ (\lambda^2+\alpha_1^2)\frac{\sigma_0^2}{2\pi} + \lambda^2\frac{\sigma_1^2}{2\pi} &= (1+\theta_1^2\lambda^2)\frac{\sigma_u^2}{2\pi}, \end{aligned}$$

where $i=\sqrt{-1}$ and $-\infty < \lambda < \infty$ denotes frequency. Equality can be achieved for all λ by equating the coefficients on the powers in λ. For λ^0 we have

$$\sigma_u^2 = \alpha_1^2\sigma_0^2,$$

while for λ^2 we have

$$\theta_1 = \sqrt{\frac{\sigma_0^2+\sigma_1^2}{\sigma_u^2}} = \sqrt{\frac{\sigma_0^2+\sigma_1^2}{\alpha_1^2\sigma_0^2}}.$$

References

Aït-Sahalia, Y., 2002. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica* 70, 232–262.

Aït-Sahalia, Y., 2007. Estimating continuous-time models using discretely sampled data. In: *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society*, Blundell, R., Persson, T., Newey, W.K. (Eds.). Cambridge University Press.

Aït-Sahalia, Y., 2008. Closed-form likelihood expansions for multivariate diffusions. *Annals of Statistics* 36, 906–937.

Aït-Sahalia, Y., Jacod, J., 2014. *High-Frequency Financial Econometrics.* Princeton University Press.

Aït-Sahalia, Y., Hansen, L.P., (Eds.), 2010. *Handbook of Financial Econometrics, Volumes 1 and 2.* Elsevier.

Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, Th.V., (Eds.), 2009. *Handbook of Financial Time Series.* Springer.

Andresen, A., Benth, F.E., Koekebakker, S., Zakamulin, V., 2014. The CARMA interest rate model. *International Journal of Theoretical and Applied Finance* 17, 1–27.

Basak S., Croitoru, B., 2000. Equilibrium mispricing in a capital market with portfolio constraints *Review of Financial Studies* 13, 715–748.

Benth, F.E., Šaltytė Benth, J., 2009. Dynamic pricing of wind futures. *Energy Economics* 31, 16-24.

Benth, F.E., Koekebakker, S., Zakamulin, V., 2008. The CARMA interest rate model. Unpublished manuscript.

Bergstrom, A.R., 1983. Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica* 51, 117–152.

Bergstrom, A.R., 1984. Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics, Vol. 2*, pp. 1145–1212, North-Holland.

Bergstrom, A.R., 1990. *Continuous Time Econometric Modelling*. Oxford University Press.

Bergstrom, A.R., Nowman, K.B., 2007. *A Continuous Time Econometric Model of the United Kingdom with Stochastic Trends*. Cambridge University Press.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637-654.

Brennan, M.J., Schwartz, E.S., 1979. A continuous time approach to the pricing of bonds. *Journal of Banking and Finance* 3, 133–155.

Brockwell, P.J., 2001. Lévy-driven continuous time ARMA processes. *Annals of the Institute of Statistical Mathematics* 53, 113-124.

Brockwell, P.J., 2004. Representations of continuous-time ARMA processes. *Journal of Applied Probability* 41, 375-382.

Brockwell, P.J., 2009. Lévy-driven continuous time ARMA processes. In: *Handbook of Financial Time Series*, Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, Th.V., (Eds.). Springer.

Brockwell, P.J., Marquardt, T., 2005. Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statistica Sinica* 15, 477-494.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The Econometrics of Financial Markets*. Princeton University Press.

Campbell, J.Y., Kyle, A.S., 1993. Smart money, noise trading and stock prices. *Review of Economic Studies* 60, 1–34

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. *Journal of Political Economy* 95, 1062–1088.

Chambers, M.J., 1991. Forecasting discrete stock and flow data generated by a second order continuous time system. *Computers and Mathematics with Applications* 22, 107–114.

Chambers, M.J., 2015. The estimation of continuous time models with mixed frequency data. *Journal of Econometrics* (forthcoming).

Chambers, M.J., Thornton, M.A., 2012. Discrete time representations of continuous time ARMA processs. *Econometric Theory* 28, 219–238.

Chambers, M.J., Thornton, M.A., 2016. Misspecification testing in continuous time models. In preparation.

de los Rios, A., Sentana, E., 2011. Testing uncovered interest parity: a continuous-time approach. *International Economic Review* 52, 1215–1251.

Florens-Zmirou, D., 1989. Approximate discrete-time schemes for statistics of diffusion processes. *Statistics* 20, 547–557.

Harrison, J.M., Pitbladdo, R., Schaefer, S.M., 1984. Continuous price processes in frictionless markets have infinite variation. *The Journal of Business* 57, 353–365

Harvey, A.C., Stock, J.H., 1985. The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1, 97–117.

Harvey, A.C., Stock, J.H., 1988. Continuous time autoregressive models with common stochastic trends. *Journal of Economic Dynamics and Control* 12, 365–384.

Huang, C.F., 1987. An intertemporal general equilibrium asset pricing model: the case of diffusion information. *Econometrica* 55, 117-142.

Jarrow, R., Protter, P., 2005. Large traders, hidden arbitrage, and complete markets. *Journal of Banking & Finance* 29, 2803-2820.

Jegadeeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48, 65–91.

Jewitt, G., McCrorie, J.R., 2005. Computing estimates of continuous time macroeconometric models on the basis of discrete data. *Computational Statistics and Data Analysis* 49, 397–416.

Koijen, R.S.J., Rodriguez, J.C., Sbuelz, A., 2009. Momentum and mean reversion in strategic asset allocation. *Management Science* 55, 1199–1213.

Liu, J., Longsaff, F.A., 2000. Losing money on arbitrage: optimal dynamic portfolio choice in markets with arbitrage opportunities. *Review of Financial Studies* 17, 611–641.

Lo, A.W., 1988. Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory* 4, 231-247

Merton, R.C., 1969. Lifetime portfolio selection under uncertainty: the continuous time case. *Review of Economics and Statistics* 51, 247-257.

Merton, R.C., 1973. Theory of rational option pricing. *Bell Journal of Economics* 4, 141-183.

Nowman, K.B., 1997. Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 52, 1695–1706.

Nowman, K.B., 1998. Continuous-time short term interest rate models. *Applied Financial*

Economics 8, 401–407.

Phillips, P.C.B., 1991. Error correction and long-run equilibrium in continuous time. *Econometrica* 59, 967–980.

Phillips, P.C.B., Yu, J., 2009. Maximum likelihood and gaussian estimation of continuous time models in finance. In: *Handbook of Financial Time Series*, Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, Th.V., (Eds.). Springer.

Priestley, M.B., 1981. *Spectral Analysis and Time Series*. Academic Press.

Robinson, P.M., 1976. Fourier estimation of continuous time models. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 215–266. North-Holland.

Robinson, P.M., 1993. Continuous-time models in econometrics: closed and open systems, stocks and flows. In: *Models, Methods, and Applications of Econometrics: Essays in Honor of A.R. Bergstrom*, Phillips, P.C.B (Ed.). Blackwell.

Sangvinatsos, A., Wachter, J.A., 2005. Does the failure of the expectation hypothesis matter for long-term investors? *Journal of Finance* 60, 179–230.

Thornton, M.A., Chambers, M.J., 2016. Continuous time ARMA processes: discrete time representation and likelihood evaluation. In preparation.

van Loan, C.F., 1978. Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, AC-23, 395–404.

Vasicek, O., 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177–188.

Wang, X., Phillips, P.C.B., Yu, J., 2011. Bias in estimating multivariate and univariate diffusions. *Journal of Econometrics* 161, 228–245.

Willard, G.A., Dybvig, P.H., 2000. Empty Promises and Arbitrage. *Review of Financial Studies* 12, 807–834.

Zadrozny, P., 1988. Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies. *Econometric Theory* 4, 108-124.

Table 1
Estimates of CARMA models for short-term interest rate at different sampling frequencies

Frequency:	Weekly	Monthly	Quarterly
CARMA(1, 0)			
a_0	0.0026	0.0029	0.0027
	(0.0023)	(0.0023)	(0.0024)
A_0	−0.0309	−0.0344	−0.0351
	(0.0249)	(0.0241)	(0.0249)
σ_u	0.0096	0.0102	0.0104
	(0.0002)	(0.0004)	(0.0007)
$\log L$	7371.5142	1370.2543	383.4108
S_{12}	0.1356	0.0797	0.4788
CARMA(2, 0)			
a_0	0.3915	0.1102	0.0204
	(0.1364)	(0.1782)	(0.0190)
A_0	−4.6676	−1.3081	−0.2590
	(0.5164)	(1.5596)	(0.2184)
A_1	−137.0820	−34.5267	−6.1614
	(18.6358)	(26.7567)	(4.4423)
σ_u	1.3806	0.3676	0.0695
	(0.1786)	(0.2691)	(0.0451)
$\log L$	7374.0029	1370.7351	384.0340
S_{12}	0.3319	0.0908	0.5812
CARMA(2, 1)			
a_0	0.0009	0.0007	0.0009
	(0.0009)	(0.0009)	(0.0011)
A_0	−0.0125	−0.0103	−0.0137
	(0.0080)	(0.0077)	(0.0091)
A_1	−0.2040	−0.1970	−0.2539
	(0.0619)	(0.0662)	(0.0870)
θ	4.7961	6.4076	4.7971
	(2.2209)	(3.5194)	(2.4909)
σ_u	0.0020	0.0016	0.0021
	(0.0009)	(0.0009)	(0.0011)
$\log L$	7376.8255	1376.1754	385.4889
S_{12}	0.3528	0.2454	0.6717
$LR(\theta = 0)$	0.0175	0.0001	0.1131

Table 2
Estimates of CARMA models for short-term interest rate using Euler approximation

	Weekly	Monthly	Quarterly
CARMA$(1,0)$			
a_0	0.0026	0.0029	0.0027
	(0.0033)	(0.0042)	(0.0042)
A_0	−0.0308	−0.0342	−0.0345
	(0.0272)	(0.0333)	(0.0356)
σ_u	0.0096	0.0101	0.0102
	(0.0001)	(0.0002)	(0.0004)
$\log L$	7371.5142	1370.2543	383.4108
S_{12}	0.1364	0.0821	0.4962
CARMA$(2,0)$			
a_0	0.0364	0.0091	0.0029
	(0.0438)	(0.0125)	(0.0043)
A_0	−0.4334	−0.1085	−0.0376
	(0.3577)	(0.1000)	(0.0369)
A_1	−12.4950	−2.8344	−0.9240
	(0.2233)	(0.1220)	(0.0862)
σ_u	0.1263	0.0303	0.0101
	(0.0007)	(0.0006)	(0.0004)
$\log L$	7374.0526	1370.8145	383.7957
S_{12}	0.3381	0.0952	0.6269
CARMA$(2,1)$			
a_0	0.0114	0.0010	0.0010
	(0.0130)	(0.0012)	(0.0012)
A_0	−0.1342	−0.0123	−0.0148
	(0.1146)	(0.0100)	(0.0098)
A_1	−3.0995	−0.1811	−0.2763
	(1.6847)	(0.1581)	(0.0499)
θ	0.1988	3.9126	3.1552
	(0.1356)	(3.7428)	(1.2375)
σ_u	0.0348	0.0024	0.0024
	(0.0172)	(0.0021)	(0.0007)
$\log L$	7375.1754	1373.2746	384.9630
S_{12}	0.5969	0.1951	0.4973
$LR(\theta = 0)$	0.1340	0.0265	0.1265

Table 3
Estimated roots of CARMA models for short-term interest rate

p, q	Weekly	Monthly	Quarterly
Exact discrete time representations			
$1, 0$	-0.0309	-0.0344	-0.0351
$2, 0$	-0.0341	-0.0379	-0.0423
	-137.0480	-34.4888	-6.1191
$2, 1$	-0.1020	-0.0985	-0.0783
	$\pm 0.0461i$	$\pm 0.0252i$	$\pm 0.1756i$
Euler approximations			
$1, 0$	-0.0308	-0.0342	-0.0345
$2, 0$	-0.0348	-0.0388	-0.0427
	-12.4602	-2.7956	-0.8813
$2, 1$	-0.0439	-0.0906	-0.0727
	-3.0556	$\pm 0.0641i$	-0.2036

Table 4
Estimates for the CARMA$(1,0)$ term structure model

	Weekly	Monthly	Quarterly
α	0.0848	0.0552	0.0653
	(0.0457)	(0.0325)	(0.0417)
k_1	0.0131	0.0295	0.0647
	(0.0588)	(0.0396)	(0.0435)
k_2	0.0442	0.0304	0.0322
	(0.0429)	(0.0329)	(0.0329)
$\ln p$	0.0140	0.0638	0.1206
	(0.0119)	(0.0576)	(0.1590)
q	0.0067	0.0310	0.1529
	(0.0056)	(0.0197)	(0.0658)
σ_1	0.0891	0.0619	0.0758
	(0.0021)	(0.0031)	(0.0069)
σ_2	0.0703	0.0589	0.0577
	(0.0017)	(0.0030)	(0.0052)
ρ	0.0259	-0.0115	-0.0799
	(0.0339)	(0.0804)	(0.1285)
Roots	0.0922	0.0897	0.1463
	0.0065	0.0293	0.0395
$\ln L$	4257.9418	770.7349	169.3028
S_{12}	[0.0026]	[0.0000]	[0.0929]

Standard errors in parentheses; p-values in square brackets.

Table 5
Estimates for the CARMA$(2,0)$ term structure model

	Weekly	Monthly	Quarterly
α	4.9333	0.6241	0.1281
	(4.0294)	(0.3312)	(0.0936)
k_1	-2.3622	-0.6682	0.0718
	(2.5592)	(1.4070)	(1.5806)
k_2	2.7039	0.8681	0.0139
	(2.6179)	(2.0166)	(0.1764)
σ_1	6.2269	0.6601	0.1163
	(0.3095)	(0.1021)	(0.0417)
σ_2	5.2959	2.2663	0.7444
	(0.7351)	(6.6858)	(14.4853)
ρ	0.0343	0.0219	0.1827
	(0.0420)	(0.0718)	(1.7910)
γ_1	60.0521	8.7896	0.8541
	(3.6327)	(1.8480)	(0.3523)
γ_2	65.0928	36.6878	12.0768
	(11.0595)	(114.5287)	(246.3730)
Roots	-0.0036	-0.0043	-0.0074
	-0.1203	-0.0912	-0.1853
	-59.9692	-8.7179	-0.6626
	-65.0518	-36.6641	-12.0757
$\ln L$	4218.9718	774.9704	178.9368
S_{12}	[0.0000]	[0.0552]	[0.9742]

Standard errors in parentheses; p-values in square brackets.

Table 6
Estimates for the CARMA$(2,1)$ term structure model

	Weekly	Monthly	Quarterly
α	3.0501	0.0894	0.0539
	(6.4713)	(0.0448)	(0.0336)
k_1	−0.7162	−0.1666	−0.0715
	(5.2446)	(2.6682)	(0.1630)
k_2	1.1002	0.2399	0.0994
	(7.5515)	(3.8589)	(0.1681)
σ_1	3.3465	0.0620	0.0482
	(6.8452)	(0.0231)	(0.0164)
σ_2	3.9020	0.8289	0.2616
	(24.7424)	(13.4620)	(0.1405)
ρ	−0.0309	0.0550	0.4139
	(0.0491)	(0.0714)	(0.1374)
γ_1	41.6066	0.4335	0.3362
	(83.8215)	(0.2711)	(0.1225)
γ_2	64.4861	13.3527	3.8875
	(406.9634)	(217.1427)	(2.3110)
θ_1	−0.0318	0.8079	−0.9512
	(0.0546)	(0.3243)	(0.4024)
θ_2	0.0255	0.0569	−0.1016
	(0.1139)	(1.2318)	(0.1583)
Roots	−0.0051	−0.0052	−0.0064
	−0.0854	−0.2232	−0.1777
	−41.5331	$\pm 0.2125i$	$\pm 0.1701i$
	−64.4691	−13.3348	−3.8618
$\ln L$	4266.2651	785.9869	182.3819
S_{12}	[0.6785]	[0.8970]	[0.9762]
$LR(\theta_1 = \theta_2 = 0)$	[0.0000]	[0.0000]	[0.0319]

Standard errors in parentheses; p-values in square brackets.

Table 7
Estimates of cointegrated CARMA$(1,0)$ model for stock prices and dividends

	$Ds(t)$	$Dd(t)$
a_0'	-0.0007	-0.0283
	(0.0098)	(0.0057)
α'	0.0009	0.0140
	(0.0047)	(0.0020)
β'	1.0000	-1.4079
		(0.0959)
Q'	0.0420	0.0004
	(0.0008)	(0.0005)
	-	0.0181
		(0.0003)
$\log L$		6330.3980
S_4 S_{12}	[0.0000]	[0.0000]

Table 8
Estimates of cointegrated CARMA$(2,0)$ model for stock prices and dividends

	$D^2s(t)$	$D^2d(t)$
a_0'	0.0084	-0.0960
	(0.0457)	(0.0211)
A_1'	-2.6594	-0.6519
	(0.4011)	(0.2353)
	-0.8612	-2.9394
	(1.1588)	(0.3779)
α'	-0.0021	0.0471
	(0.0220)	(0.0089)
β'	1.0000	-1.3978
		(0.0863)
Q'	0.1415	0.0379
	(0.0200)	(0.0180)
	-	0.0560
		(0.0051)
$\log L$		6426.6386
S_4 S_{12}	[0.0000]	[0.0000]

Table 9
Estimates of cointegrated CARMA$(2, 1)$ model for stock prices and dividends

	$D^2s(t)$	$D^2d(t)$
a_0'	0.0316	-0.0125
	(0.0317)	(0.0063)
A_1'	-1.8575	0.2650
	(0.4914)	(0.1434)
	0.3948	-0.2111
	(0.6330)	(0.1104)
α'	-0.0156	0.0065
	(0.0159)	(0.0029)
β'	1.0000	-1.4832
		(0.1038)
Θ'	0.2190	0.4060
	(0.1863)	(0.1742)
	0.2749	2.5502
	(1.7101)	(0.3917)
Q'	0.1009	-0.0144
	(0.0245)	(0.0058)
	-	0.0181
		(0.0008)
$\log L$		6504.4292
S_4 S_{12}	[0.1673]	[0.0919]

Table 10
Estimates of cointegrated CARMA$(1, 0)$ model for stock prices and dividends, both stocks

	$Ds(t)$	$Dd(t)$
a_0'	-0.0029	-0.0273
	(0.0100)	(0.0049)
α'	0.0019	0.0132
	(0.0047)	(0.0017)
β'	1.0000	-1.3874
		(0.0868)
Q'	0.0420	0.0013
	(0.0008)	(0.0004)
	-	0.0155
		(0.0003)
$\log L$		6212.1120
S_4 S_{12}	[0.0000]	[0.0000]

Table 11
Estimates of cointegrated CARMA$(2,0)$ model for stock prices and dividends, both stocks

	$D^2 s(t)$	$D^2 d(t)$
a_0'	0.0258	-0.0454
	(0.0317)	(0.0094)
A_1'	-2.4507	-0.2153
	(0.2178)	(0.0454)
	-0.2059	-1.3383
	(0.3547)	(0.0935)
α'	-0.0108	0.0224
	(0.0148)	(0.0036)
β'	1.0000	-1.4063
		(0.0949)
Q'	0.1307	0.0122
	(0.0084)	(0.0023)
	-	0.0301
		(0.0011)
$\log L$		6452.3617
S_4 S_{12}	[0.0000]	[0.0000]

Table 12
Estimates of cointegrated CARMA(2, 1) model for stock prices and dividends, both stocks

	$D^2s(t)$	$D^2d(t)$
a_0'	0.0465	-0.0149
	(0.0284)	(0.0037)
A_1'	-1.9047	0.0295
	(0.5414)	(0.0777)
	0.8903	-0.3028
	(0.4268)	(0.0490)
α'	-0.0231	0.0078
	(0.0134)	(0.0014)
β'	1.0000	-1.4692
		(0.1010)
Θ'	0.1625	0.0345
	(0.1672)	(0.0281)
	1.2289	1.0286
	(0.4932)	(0.1239)
Q'	0.1039	-0.0037
	(0.0265)	(0.0040)
	-	-0.0095
		(0.0009)
$\log L$		6512.1382
S_4 S_{12}	[0.9999]	[0.0465]

Table 13
Intercept and short-run dynamics of discrete representations of CARMA (2, 1) models

	f_0	a	F_1	
stock-flow treatment				
stock price	0.0118	-0.0058	-0.2134	-0.1443
dividend	-0.0124	0.0064	-0.3013	-0.7959
stock-stock treatment				
stock price	0.0173	-0.0072	-0.1523	-0.3305
dividend	-0.0125	0.0068	-0.0106	-0.7450

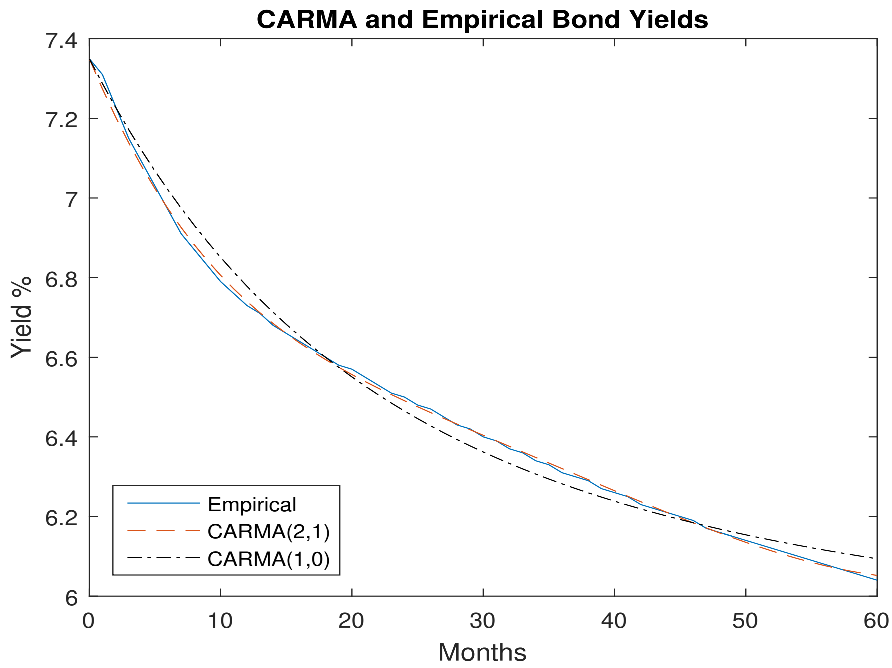

Figure 1. Empirical and CARMA Bond Yields, 31 March 1998

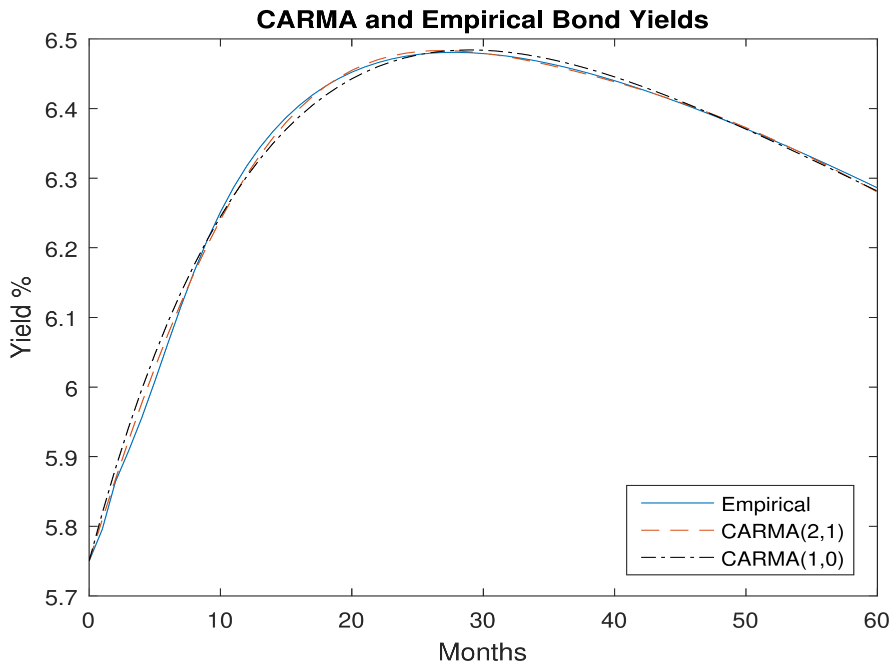

Figure 2. Empirical and CARMA Bond Yields, 31 January 2000

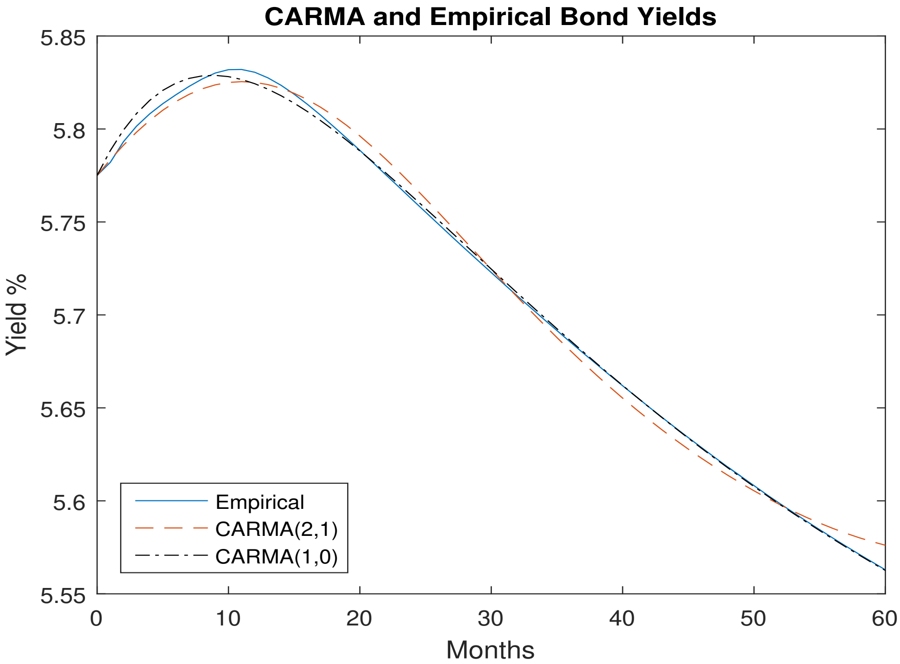

Figure 3. Empirical and CARMA Bond Yields, 30 June 2007

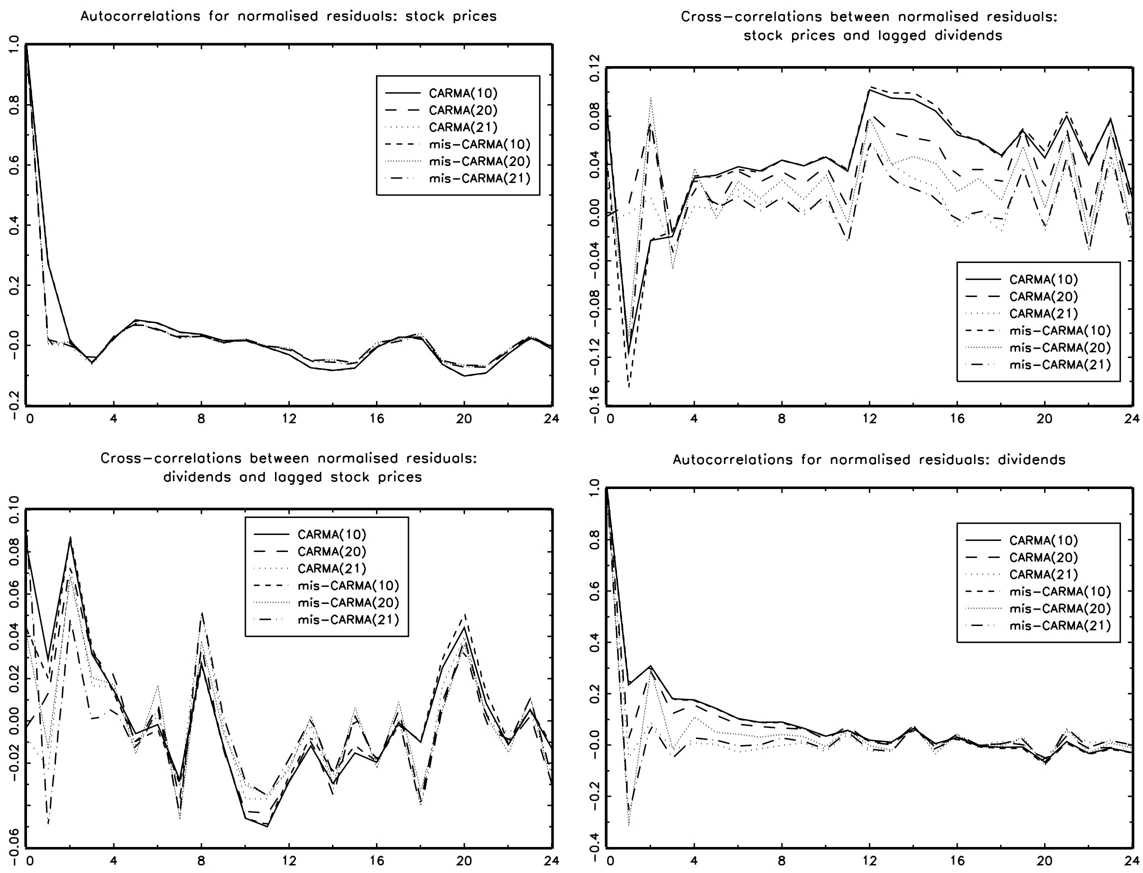

Figure 4. Autocorrelations and Cross-correlations of the normalised residual vector e_t for the Stock price and Dividends model